Chemistry Central Journal

Poster presentation

Open Access

Molecular descriptors based on entropy and the full topological neighborhood of all atoms

K Varmuza*[1], M Dehmer[2] and S Borgert[3]

Address: [1]Vienna University of Technology, Institute of Chemical Engineering, Getreidemarkt 9/166, A-1060 Vienna, Austria, [2]Coimbra, Portugal and [3]Darmstadt, Germany

* Corresponding author

from 4th German Conference on Chemoinformatics
Goslar, Germany. 9–11 November 2008

Published: 5 June 2009

Chemistry Central Journal 2009, **3**(Suppl 1):P38 doi:10.1186/1752-153X-3-S1-P38

This abstract is available from: http://www.journal.chemistrycentral.com/content/3/S1/P38

A new family of topological information indices based on the full neighbourhood of all atoms is presented. In previous definitions, mostly partitions of atoms have been used for this purpose. We consider each atom of a molecular structure as a subsystem. For each atom the complete neighbourhood is characterized by an information functional, based on the number of atoms in all spheres around the atom. An appropriate weighting scheme summarizes the number of atoms in the different spheres, resulting in a characteristic property of the atom. The properties of all atoms are normalized to a sum of one (a probability-like measure) from which the information entropy is calculated. The entropy is scaled by the number of atoms in the structure to give a molecular descriptor E. In the current version only skeletons of the chemical structures are considered, with all atoms equal and all bonds equal.

The topological information index E has been calculated for some complete sets of isomers (trees, one ring, two rings) with up to 12 atoms. Furthermore, a selection of 4,000 structures from a mass spectral data base has been available. Index E is compared with about 200 other molecular descriptors by applying multivariate data analysis.

The new index characterizes the diversity of the atoms in terms of neighbourhood, and thereby a special type of structural complexity and inner symmetry. If the neighbourhoods of the atoms are very different among each other, E is large with typical values around 10. If the atoms are topologically identical, E is zero, for instance for unsubstituted rings. Thus we generalized the classical information indices from chemical graph theory because the new descriptor type is parameterized and allows the incorporation of different information functionals.